JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1986, 19, 73–77 NUMBER 1 (SPRING 1986)

# METHODS OF TIME SAMPLING: A REAPPRAISAL OF MOMENTARY TIME SAMPLING AND PARTIAL INTERVAL RECORDING

ALEX HARROP AND MICHAEL DANIELS

LIVERPOOL POLYTECHNIC

We compared the accuracy of momentary time sampling (MTS) and partial interval recording (PIR) in estimating both absolute behavioral levels and relative change. A computer randomly generated runs of pseudobehavior varying in duration and rate and simulated MTS and PIR of each run. Results indicated that when estimating absolute behavioral levels, duration rather than rate should be used as the dependent measure, and MTS is more accurate than PIR. In contrast, PIR is the more sensitive method for detecting relative changes in behavioral levels, although, at high rates, PIR tends to underestimate the degree of change.

DESCRIPTORS: observation methods, momentary time sampling, partial interval recording, measurement error

---

Time sampling is a procedure that, although suffering from inherent limitations (Johnston & Pennypacker, 1980), is widely used in applied behavior analysis. The accuracy of time-sampling methods has been investigated by several researchers, in particular by Repp, Roberts, Slack, Repp, and Berkler (1976), and by Powell, Martindale, Kulp, Martindale, and Bauman (1977). The two methods that were compared in both investigations may be defined as (a) momentary time sampling (MTS), in which a response is scored if it occurs exactly at a predetermined moment, and (b) partial interval recording (PIR), in which an observation interval is scored if a response occurs during any part of the interval.

Repp et al. (1976) examined the accuracy of MTS and PIR in assessing rate of responding. They found MTS to be extremely inaccurate for all conditions investigated. PIR was found to be accurate for low and medium rates of responding, but to underestimate high-rate responding. Powell et al. (1977) examined the accuracy of MTS and PIR in assessing duration of responding. They found MTS to be superior to PIR, which overestimated duration. The work of Repp et al. therefore suggests that PIR is the better method for measuring rate of responding, whereas that of Powell et al. suggests that MTS is better for measuring duration.

An examination of these studies indicates, however, that neither adequately equated the conditions under which MTS and PIR were compared, and neither explored the possible independent influence of behavioral parameters (e.g., duration of behavioral episodes and rate of responding) on the accuracy of the techniques. Furthermore, the analyses applied to the data were, in both cases, limited to a comparison of accuracy in estimating absolute behavioral levels. Although such estimation is, in many situations, important (e.g., in determining whether an intervention is necessary), an equally important consideration, particularly in intervention research, is accuracy in the estimation of relative changes in behavioral level.

In view of the limitations in the studies by Repp et al. (1976) and Powell et al. (1977), we decided to examine further the accuracy of comparable MTS and PIR procedures in estimating both absolute behavioral levels and relative changes, for various durations and rates of behavior. We hope that our results may help to provide researchers and practitioners with clearer guidelines on the relative merits of the two techniques and on the conditions in which one method may be preferred.

Requests for reprints should be sent to Alex Harrop, Section of Psychology, Liverpool Polytechnic, C. F. Mott Campus, Liverpool Road, Prescot, Merseyside L34 1NP, United Kingdom. Michael Daniels is at the same address.

Table 1
Mean Percentage Error of Absolute Estimation for Momentary Time Sampling (MTS) and Partial Interval Recording (PIR)

| Emitted duration (s) | Low to medium rates | | Medium to high rates | |
|---|---|---|---|---|
| | MTS | PIR | MTS | PIR |
| | Estimated rate | | | |
| 1 | 5.5 | −8.5** | 3.9 | −36.5** |
| 5 | 387.5** | 32.0** | 396.7** | 2.7 |
| 10 | 900.5** | 82.5** | 891.9** | 53.1** |
| 20 | 1,871.6** | 181.2** | 1,883.4** | 151.2** |
| | Estimated duration | | | |
| 1 | 5.5 | 815.2** | 3.9 | 535.0** |
| 5 | −2.5 | 164.0** | −0.7 | 105.4** |
| 10 | 0.1 | 82.5** | −0.8 | 53.1** |
| 20 | −1.4 | 40.6** | −0.8* | 25.6** |

* $p < .01$.
** $p < .0001$.

## METHOD

### *Design*

To permit a fair comparison between MTS and PIR, we attempted to equate the costs, and the demands on the observer, of these techniques. A "time base" of 15 s defined the time between the start of successive observations. For MTS, this meant that behavior was observed for 1 s every 15 s. For PIR, the behavior was observed for a 10-s observation interval, with a nonobservation time (used in practice for recording) of 5 s. We chose a total sampling period of 1 hr for both methods, representing a period for which a human observer might realistically be expected to remain on task. For each method, therefore, 240 observations occurred in the session.

The emitted behavior was regulated into four constant durations (1, 5, 10, and 20 s), representing behaviors that were almost instantaneous, one-half, equal to, and twice the observation interval used for PIR. The use of a 20-s duration also permitted a behavior to span two observation instances or intervals. We chose constant duration behaviors because, although unlikely in practice, these permit a controlled parametric investigation.

The rates of emission were controlled by setting the probability of onset of a behavior ($p$) into two sequences: (a) "low to medium" rates, where $p$ = 1:180, 1:90, 1:60, 1:45, 1:36, and 1:30, representing expected frequencies of 20, 40, 60, 80, 100, and 120 per hr for behaviors of 1-s duration, and (b) "medium to high" rates, where $p$ = 1:30, 1:15, 1:10, 1:7.5, 1:6, and 1:5, representing expected frequencies of 120, 240, 360, 480, 600, and 720 per hr for behaviors of 1-s duration.

### *Computer Simulation*

We chose computer simulation for our study because it offers the advantages of speed, accuracy, and precise parametric control. Simulation was based on the assumption that, when using human observers, 1 s represents the shortest time in which a behavior may occur and be observed. Accordingly, time was simulated in terms of successive discrete "moments" (considered, notionally, as seconds) in which a behavior may either occur or not occur.

For each combination of emitted duration and probability of onset, a BBC Model B microcomputer generated 20 runs of pseudobehavior, producing 880 separate behavioral records (20 runs × 4 durations × 11 probabilities). To produce these records, a Basic procedure randomly generated the pseudobehaviors into a single-dimension string array of 3,620 elements. This size covered the sampling period (3,600 s), preceded by a further 20 s that permitted behaviors to be initiated prior to the commencement of observation. The Basic random number function controlled the probability of onset of behavior. We appreciated inadequacies of the Basic function as a source of random numbers, but judged that this would not materially affect the validity of the simulation procedure. Once a behavior was initiated, another occurrence was not permitted until the first behavior was completed. This ensured that all behaviors were of constant duration, although successive occurrences could follow without pause.

Following the generation of each behavioral record, a further procedure sampled the array every 15 s, beginning with the 21st element, simulating

Table 2

Indices of Sensitivity and Linearity for Momentary Time Sampling (MTS) and Partial Interval Recording (PIR)

| Emitted duration (s) | Low to medium rates | | | | Medium to high rates | | | |
|---|---|---|---|---|---|---|---|---|
| | Sensitivity | | Linearity[a] | | Sensitivity | | Linearity[a] | |
| | MTS | PIR | MTS | PIR | MTS | PIR | MTS | PIR |
| 1 | 0.049 | 0.538 | 0.993 | 0.999 | 0.037 | 0.371 | 0.998 | 0.949** |
| 5 | 0.397 | 0.863 | 0.995 | 0.999 | 0.340 | 0.729 | 0.997 | 0.969** |
| 10 | 0.611 | 0.866 | 0.999 | 0.999* | 0.390 | 0.688 | 0.999 | 0.977** |
| 20 | 0.846 | 0.949 | 0.999 | 1.000 | 0.579 | 0.800 | 0.999 | 0.987** |

[a] Significance levels are for curvilinearity.
* $p < .01$.
** $p < .0001$.

both MTS and PIR. For MTS, the procedure counted the number of observations that recorded behavior. For PIR, the procedure counted the number of intervals during which, at any time, behavior occurred.

## RESULTS AND DISCUSSION

### *Estimating Absolute Levels*

To examine the error that may be produced by MTS and PIR when estimating absolute behavioral levels, actual rate and total duration of behavior were compared with estimates derived from the recorded data. For rate, the actual frequency with which behavior occurred in the sampling period was compared with the frequency estimated per hour of observation time (recorded frequency × 3,600/seconds of observation). For duration, the actual proportion of the sampling period during which behavior occurred was compared with the proportion of observations that were scored.

Results indicate that error of estimation is a function of dependent measure, sampling method, emitted duration, and, for PIR (but not MTS), behavioral rate. Table 1 presents the mean percentage errors of estimation under the various conditions. To test for the presence of systematic overestimation or underestimation, the binomial test was applied to the estimates, following the procedure used by Brulle and Repp (1984).

When estimating absolute rate, considerable systematic error is, in general, produced by both sampling methods, with rates being progressively overestimated as emitted duration increases. For PIR, overestimation also increases with lower behavioral rates. With short duration behaviors, however, particularly if they are also of higher rates, PIR underestimates rate. It is pertinent to note here that Repp et al. (1976), who found that PIR was accurate in estimating rates or, with high rates, produced underestimation, used pseudobehaviors of very short duration (0.035 s). Our results demonstrate clearly the problems of using rate as the dependent measure with time-sampling procedures; problems due to the impossibility, given only information that, for example, 200 out of 240 observations are scored, of determining whether this represents behavior of high frequency and short duration or of low frequency and long duration.

When estimating absolute duration, MTS appears not to introduce systematic error, whereas PIR produces overestimation that increases with shorter emitted durations and lower rates. These results are consistent with those of Powell et al. (1977), who concluded that MTS offers clear advantages over PIR when duration is the measure of interest.

### *Estimating Relative Changes*

To examine the accuracy of MTS and PIR in estimating relative changes in behavioral level, the criterion adopted was the regression of recorded on actual rates. Because emitted duration is constant

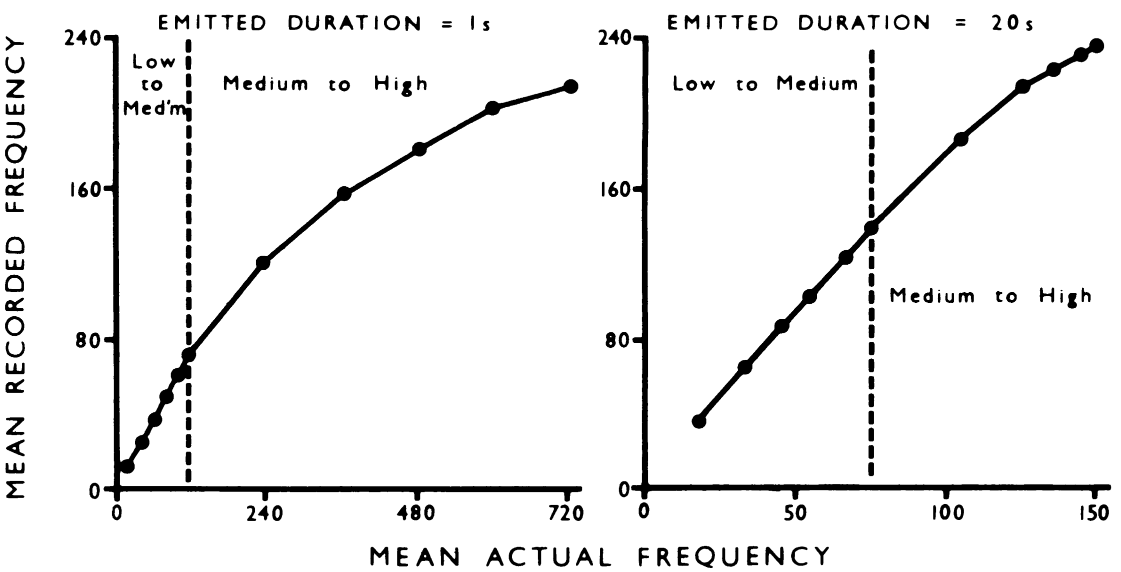


Figure 1. Mean recorded frequency obtained with partial interval recording as a function of mean actual frequency.

in each condition, regression analyses for rate are equivalent to those for total duration. Analysis of variance of regression was carried out on the data from each of the two probability sequences for each sampling method and emitted duration. An index of *sensitivity* is provided by the proportion of variance *within* probability levels accounted for by pooled regression. Sensitivity is a measure of the ability of a sampling method to reflect accurately small changes in the actual rate of behavior (randomly occurring within each probability level). An index of *linearity* is provided by the proportion of variance *between* probability levels accounted for by linear regression. Linearity is a measure of the degree to which a graph of the average recorded rates across probability levels would be similar in shape to a graph of the average actual rates. Table 2 presents the indices of sensitivity and linearity for both methods of recording in the various conditions.

With both sampling methods, sensitivity increases with lower rates and longer emitted durations. More importantly, in all conditions, sensitivity is significantly ($p < .005$) and substantially greater with PIR than with MTS. For low to medium rates, linearity with both sampling methods is high for all emitted durations. For medium to high rates, linearity with MTS is high for all emitted durations, whereas with PIR it is lower because of a significant curvilinear component. To examine curvilinearity, graphs were drawn to show the relationship between mean recorded frequency and mean actual frequency for each emitted duration (for examples, see Figure 1). Graphs indicate the presence of systematic error with PIR (attenuated at longer emitted durations) that is in the direction of underestimating change with high rates of behavior.

## *Conclusions*

Although caution must be exercised in extrapolating our results beyond the range of values we selected for the sampling and behavioral parameters, several general conclusions are indicated. MTS (but not PIR) provides accurate average estimates of absolute duration. Estimates of absolute rate are inaccurate with both methods. PIR is more sensitive than MTS in detecting relative changes in behavioral level (rate or total duration), but PIR underestimates the magnitude of change with high-rate behaviors, particularly if they are also of short duration. Although MTS is the less sensitive method, it appears not to suffer from systematic error in estimating relative change. Because, however, the systematic error produced by PIR is always in the direction of providing a conservative estimate of change, researchers and practitioners may consider that this error is a price worth paying for the greater sensitivity of the method.

## REFERENCES

Brulle, A. R., & Repp, A. C. (1984). An investigation of the accuracy of momentary time sampling procedures with time series data. *British Journal of Psychology,* **75,** 481–485.

Johnston, J. J., & Pennypacker, H. S. (1980). *Strategies and tactics of human behavioral research.* Hillsdale, NJ: Lawrence Erlbaum.

Powell, J., Martindale, B., Kulp, S., Martindale, A., & Bauman, R. (1977). Taking a closer look: Time sampling and measurement error. *Journal of Applied Behavior Analysis,* **10,** 325–332.

Repp, A. C., Roberts, D. M., Slack, D. J., Repp, C. F., & Berkler, M. S. (1976). A comparison of frequency, interval and time-sampling methods of data collection. *Journal of Applied Behavior Analysis,* **9,** 501–508.

*Received April 8, 1985*
*Final acceptance October 29, 1985*